De novo designed protein-interaction modules for in-cell applications

AUTHORS:

Caitlin L. Edgell [a,b], Abigail J. Smith [b,c], Joseph L. Beesley [a], Nigel J. Savery [b,c,1] and Derek N. Woolfson [a,b,c,1]

[a] School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom

[b] School of Biochemistry, University of Bristol, Bristol, BS8 1TD, United Kingdom

[c] BrisSynBio, Life Sciences Building, University of Bristol, Bristol, BS8 1TQ, United Kingdom

AUTHOR INFORMATION:

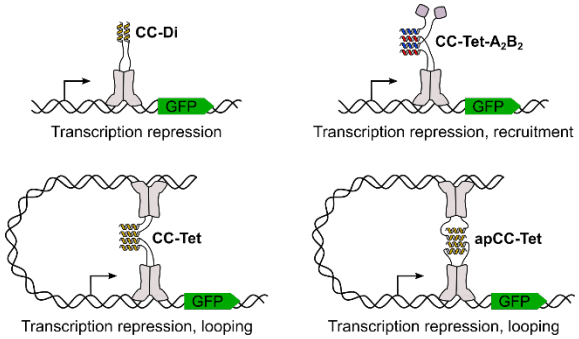
Corresponding authors:

[1] To whom correspondence should be addressed:

Nigel J. Savery, N.J.Savery@bristol.ac.uk, School of Biochemistry, University of Bristol, Bristol, BS8 1TD, United Kingdom.

Derek N. Woolfson, D.N.Woolfson@bristol.ac.uk, School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom.

TOC FIGURE:



ABSTRACT:

Protein-protein interactions control a wide variety of natural biological processes. α-Helical coiled coils frequently mediate such protein-protein interactions. Due to the relative simplicity of their sequences and structures and the ease with which properties such as strength and specificity of interaction can be controlled, coiled coils can be designed *de novo* to deliver a variety of non-natural protein-protein-interaction domains. Herein, several *de novo* designed coiled coils are tested for their ability to mediate protein-protein interactions in *Escherichia coli* cells. The set includes a parallel homodimer, a parallel homotetramer, an antiparallel homotetramer and a newly designed heterotetramer, all of which have been characterized *in vitro* by biophysical and structural methods. Using a transcription repression assay based on reconstituting the Lac repressor, we find that the modules behave as designed in the cellular environment. Each design imparts a different property to the resulting Lac repressor-coiled coil complexes, resulting in the benefit of being able to reconfigure the system in multiple ways. Modification of the system also allows the interactions to be controlled: assembly can be tuned by controlling the expression of the constituent components; and complexes can be disrupted through helix sequestration. The small and straightforward *de novo* designed components that we deliver are highly versatile and have considerable potential as protein-protein-interaction domains in synthetic biology where proteins must be assembled in highly specific ways. The relative simplicity of the designs makes them amenable to future modifications to introduce finer control over their assembly and to adapt them for different contexts.

KEYWORDS:

De novo protein design, coiled coil, transcriptional control, Lac repressor, artificial transcription factor.

ABBREVIATIONS:

bp, base pairs; CC, coiled coil; *E. coli*, *Escherichia coli*; PPI, protein-protein interaction; RNAP, RNA polymerase; SUMO, small ubiquitin-related modifier.

Protein-protein interactions (PPIs) are involved in most biological processes. As such, modulation of these interactions allows control over the behavior of cells and whole organisms.[1] Thus, the introduction of exogenous PPIs into cells has broad applications in synthetic biology where they may be used, for example, in the assembly of nanostructures or artificial enzyme cascades, or for regulating the sub-cellular localization of proteins and protein assemblies of interest.[2-5]

Transcription regulation is one process where PPIs play well-understood roles, and where the potential for controlling cellular and organismal behavior is high. Here, transcription factors allow a biological system to control the time, location and level at which its genes are expressed. This careful spatiotemporal regulation facilitates many complex natural processes including differentiation, multicellularity, reproduction and environmental sensing. Such events can also be exploited by synthetic biologists to introduce new behaviors into organisms through the design of synthetic gene regulatory networks.[6]

In their simplest forms, transcription factors achieve transcription activation by recruiting and stabilizing RNA polymerase (RNAP) at a promoter, or repression by preventing RNAP from accessing a promoter. In most cases, transcription factors bind DNA as oligomers, and, in many cases, they interact with other proteins to exert their function. Thus, PPIs are usually required for both repression and activation.[7,8] In principle, therefore, combining PPI domains with protein-DNA interaction domains allows the design of artificial transcription factors that may be used to modulate cellular and organismal behavior with high specificity. Such systems also provide a test bed for probing potential PPIs for use in other synthetic-biology applications.

PPI domains for synthetic biology can be sourced from natural systems. For example, the Gal4 dimerization domain-Gal11P interaction has been used to mediate interactions between DNA-binding proteins and between bacterial RNAP subunits and DNA-binding proteins,[9-11] and the λ cI oligomerization domain has been used to make highly cooperative transcriptional repressors when fused to various DNA-binding domains.[12] A split maltose-binding protein, the FK506-binding protein, and a dimerizing single-domain antibody have also been proposed as inducible PPIs.[13-15] However, the reuse of natural proteins can make it difficult to find PPI domains that are orthogonal to host PPIs. Also, precision protein engineering to deliver desired properties for a given application can be frustrated by the inherent complexity of natural protein domains. *De novo* protein design—that is, designing protein sequences to adopt prescribed structures and perform desired functions from scratch—offers an alternative source of PPIs. *De novo* PPIs can be designed to be relatively simple and their properties should be better understood and more controllable than those of their natural counterparts. However, it is generally challenging to design PPIs given the complex and often diffuse nature of protein-protein interfaces.[16] Nonetheless, there is one class of PPI for which the interactions can be both predicted and designed from first principles; namely, the α-helical coiled coil (CC).

CCs are widespread protein-folding motifs often found in natural PPIs, including in transcriptional regulators.[17-19] CC sequences usually consist of repeating seven-residue motifs called heptads, which are labelled ***abcdefg*** and mainly have hydrophobic (***h***) residues at the ***a*** and ***d*** positions. These ***hxxhxxx*** patterns direct the folding of amphipathic α helices, two or more of which wrap around each other to bury their hydrophobic surfaces and form rope-like superhelices.[20-22] The relative simplicity of CCs has enabled the elucidation of sequence-to-structure relationships that can be used as rules to guide the *de novo* design of sequences to adopt specific CC structures.[20] Designable characteristics include stability,[23] oligomeric state,[24] whether an assembly is homo- or heteromeric,[25] and helix orientation,[26,27] which are all accessible through variation of the heptad sequence repeat.

This versatility of CCs makes them useful as plug-and-play PPI domains, particularly in the context of transcriptional regulation.[28-32] Natural and designed CCs have also previously been used to reconstitute split proteins [33-35] and to co-assemble diverse proteins of interest to give multivalent antibodies,[36] oligomeric photosynthetic reaction centers,[37] cytosolic nanostructures,[2] and protease-controlled regulatory networks.[38] In these examples the designed CCs retain their original properties when fused to bulky proteins of interest, and they appear to function well with little or no optimization for the cellular environment.

Here, we test systematically a number of *de novo* designed CCs for their ability to mediate PPIs in *Escherichia coli* (*E. coli*) cells using transcription repression as the readout. The set initially consisted of designs with diverse quaternary structures that had previously been characterized extensively *in vitro* and through to X-ray crystal structures, including parallel and antiparallel homotetrameric CCs.[24,27,39] Furthermore, in order to control fully the formation of designed PPIs *in vivo*, we set out to design a system in which the assembly of functional domains depended on the expression of a second, regulatory component. Such a system requires a heteromeric CC with an oligomeric state greater than two, for example a heterotetramer. As our initial tool kit did not contain a heterotetrameric CC, we have designed a new obligate A_2B_2 heterotetramer using sequence-to-structure relationships established by us and others.[20,21] Heterotetrameric CCs have been achieved in the past by mutating naturally occurring homotetramers to convert them from one- to two-peptide systems,[40-44] and the sequences of these CCs are generally very similar to the naturally derived parents. Therefore, the design presented here represents a truly *de novo* heterotetrameric CC in that it has been designed entirely from first principles. This heterotetramer allows the design goal of a tunable transcriptional repressor in *E. coli* to be realized, showing that downstream activity can be regulated by controlling CC assembly.

All of the discussed CCs assemble as designed *in vitro* and *in vivo* and will be useful in a range of synthetic-biology applications, particularly transcriptional regulation. Our results expand the set of *de novo* designed PPIs that have been validated for use in *E. coli* and demonstrate the practicality of using established CC design rules to address unmet challenges in protein assembly in living cells.

RESULTS AND DISCUSSION:

A set of de novo homomeric coiled coils interact in E. coli

Recently, *de novo* dimeric CCs have been used to bring together natural and engineered DNA-binding domains to control transcription in *E. coli* (Figure 1a, Table 1).[23,24,28,29] For this new study, we sought to expand this set to include the parallel and antiparallel homotetramers, CC-Tet and apCC-Tet (Figures 1b and 1c, Table 1), which have been characterized thoroughly biophysically and structurally.[24,27,39] Tetramers were selected because they fold more cooperatively than dimers, and they increase the complexity of assemblies that are accessible. These are desirable features for synthetic biology that provide increased controllability over interactions and greater resistance to noise.[45] Highly cooperative transcriptional regulators have previously been implemented in synthetic gene circuits in both eukaryotes and prokaryotes.[12,46]

The tetramers were tested for their ability to mediate PPIs in *E. coli* using a transcription repression assay where the CCs assemble oligomers of a monomeric Lac repressor variant, LacI*. Wild-type (WT) LacI is a homotetramer, but effectively it is a dimer of dimers: the protomers dimerize and the tetramer is formed by the association of *C*-terminal helices from each protein chain into a four-helix bundle. Therefore, to generate monomers, both the tetramerization domain and dimer interface must be disrupted. To do this, in LacI*, the *C*-terminal tetramerization domain was removed and a Leu251→Ala substitution that weakens dimerization was introduced.[47,48] As previously performed with CC-Di,[29] CC-Tet and apCC-Tet were each fused to the *C* terminus of LacI*; *i.e.*, effectively replacing the Lac repressor's WT tetramerization domain (Figures 1d-1g).[18,49] When expressed in *E. coli* at basal levels from the P_{araBAD} promoter, both new fusion proteins repressed the reporter gene, GFP, which was expressed from the *lacUV5* promoter (Figure 1h). Thus, like CC-Di, the tetrameric CCs restored assembly of the LacI* complex and its DNA-binding activity; and both represent useful additions to the set of *de novo* CCs that have been tested inside cells.[2,3,28,29] Moreover, the constructs with the *de novo* CCs gave stronger levels of repression than a version of LacI* with its WT tetramerization domain intact, LacI*-WTtet (Figure 1h).

LacI*-CC-Tet and LacI*-apCC-Tet should assemble two DNA-binding LacI* dimers, which opens the possibility of using them to loop DNA by binding two DNA sites simultaneously, thereby increasing repression (Figures 1f and 1g).[50,51] This is not expected to occur with the LacI*-CC-Di constructs (Figure 1e). Indeed, when the repression assay was performed with a reporter plasmid containing two O_1 *lac* operators spaced 92 base pairs (bp) apart,[29] both LacI*-CC-Tet and LacI*-apCC-Tet gave stronger repression than with a single operator site, while LacI*-CC-Di did not (Figure 1h). LacI*-apCC-Tet gave an enhancement in repression of 6.5 on addition of the second operator. This is greater than the 2.2-fold enhancement seen with LacI*-WTtet and is comparable to the enhancement observed with the WT Lac repressor. Therefore, complex assembly by the *de novo* CC alone is comparable to complex assembly where the WT dimerization and tetramerization interfaces are both intact.

Design and characterization of a novel heterotetramer, CC-Tet-A_2B_2

The above homomeric CCs provide synthetic biologists with a range of options for the assembly of multi-component complexes within cells. However, some important applications cannot be met with homomeric systems. Examples include the subcellular localization or the colocalization of functionally distinct proteins. Therefore, we sought to create a system in which the assembly of a functional complex was controlled by a separate regulatory component. In order to achieve this, it was necessary to design and validate a new type of CC component: an obligate heterotetrameric CC.

We took the parallel homotetramer, CC-Tet,[24] as a starting point because the design principles for parallel CCs are better developed than for antiparallel assemblies. An obligate A_2B_2 heterotetramer, CC-Tet-A_2B_2, was derived by redesigning CC-Tet to give two oppositely charged peptides, CC-Tet-A and CC-Tet-B, with all glutamate or all lysine residues at the core-flanking *e* and *g* positions of the CC heptad repeat, respectively (Figure 2a, Table 1). These peptides were synthesized by solid-phase peptide synthesis, purified to homogeneity (Figures S1 and S2) and investigated alone and combined in equimolar concentrations.

Charge repulsion between like peptides was expected to preclude homomer formation but allow heteromer formation between the A and B chains. Consistent with this, circular dichroism (CD) spectroscopy showed that in isolation both peptides were largely unfolded in aqueous solution with mean residue ellipticity at 222 nm (MRE_{222}) values of ≈ -14000 $deg.cm^2.dmol^{-1}.res^{-1}$ (Figures 2b, S3 and S4, Table S1). Furthermore, when mixed they interacted to form a highly α-helical heteromeric species with an MRE_{222} value of ≈ -33000 $deg.cm^2.dmol^{-1}.res^{-1}$ Figures 2b, S5). This heteromer was thermally stable and did not display an unfolding transition when heated to 95 °C. Conversely, the isolated peptides were essentially entirely unfolded over the investigated temperature range (Figure 2c). The peptides and heteromer were also investigated using analytical ultracentrifugation to determine their solution-phase molecular weights. Both peptides were monomeric in isolation, while the mixture formed a tetramer (Figures 2d-2g, Tables S1-S3). Finally, to determine the stoichiometry of this heterotetramer, continuous variation experiments were performed. CD spectra were recorded at different ratios of CC-Tet-A and CC-Tet-B (Figure S6a), and the fraction helix observed was plotted against the ratio in a Job plot (Figure S6b). The highest amount of helix was observed for the 1:1 ratio of CC-Tet-A:CC-Tet-B.

Together, these biophysical data show that the designed peptides form an obligate A_2B_2 heterotetramer *in vitro* as designed, without significant off-target homomeric interactions. As it is based on CC-Tet, which has all-parallel helices, we anticipate that CC-Tet-A_2B_2 is also a parallel coiled-coil tetramer. However, we have not been able to obtain crystals for the complex, and, therefore, cannot confirm this by X-ray crystallography at present.

The CC-Tet-A_2B_2 heterotetramer functions in E. coli

To determine whether CC-Tet-A_2B_2 was suitable for use inside cells, it was tested in the transcription repression assay. This was initially done by fusing each peptide to the *C* terminus of LacI* (Figure 3a). The fusion proteins were expressed from separate plasmids at basal levels from P_{araBAD} and tested with GFP expressed from *lacUV5* as the reporter. Neither LacI*-CC-Tet-A nor LacI*-CC-Tet-B alone caused repression (Figure 3b). However, when co-expressed the fusion proteins restored repression, showing that the CC-Tet-A_2B_2 tetramer was able to direct the assembly of a heteromeric LacI* complex containing functional DNA-binding dimers. With a second O_1 *lac* operator in the reporter plasmid LacI*-CC-Tet-A_2B_2 showed a small enhancement in repression (Figure S7).

A heterotetrameric CC offers the potential to co-localize multiple different appended protein domains. For example, DNA binding domains could be assembled with transcription activation domains or DNA modifying domains to target enzyme activity to specific DNA sequences. As a proof of principle, we performed transcription repression assays in which CC-Tet-B was fused to the *C* terminus of SUMO (small ubiquitin-related modifier, Smt3) instead of LacI*. SUMO was selected as it is often used as a solubility tag for protein expression and was intended to stabilize the otherwise disordered peptide in the *E. coli* cytoplasm. In this instance, the heterotetrameric interaction was expected to produce a complex containing a single DNA-binding LacI* dimer and two discrete SUMO monomers linked *via* the heterotetrameric CC

(Figure 3c). When LacI*-CC-Tet-A and SUMO-CC-Tet-B were co-expressed transcription repression was also observed, albeit at a slightly lower level than with the LacI*-CC-Tet-A_2B_2 complex (Figure 3d). This demonstrates that the heterotetramer can be used to co-localize distinct protein domains inside cells and in the future SUMO could be replaced with more functionally interesting proteins. Furthermore, it shows that the CC sequences do not need to be fused to very large partner proteins in order to function within cells, laying the foundations for a minimized regulatory CC construct.

Dynamic transcriptional control using de novo coiled coils

All of the CCs described here have the potential to display dynamic assembly behavior through control of their expression at the transcriptional level. The heterotetramer is particularly amenable to this type of control because each component may be manipulated independently. To this end, we designed a heterotetramer-based regulatory system wherein protein domains of interest fused to one type of CC peptide would co-assemble only in the presence of a second, regulatory peptide, the expression of which could be controlled independently.

CC-Tet-A_2B_2 was used to make a tunable repression system where the level of transcription repression by constitutively expressed DNA-binding proteins could be altered by modulating the expression of a regulatory component with no DNA-binding activity (Figure 4a). To test this, the transcription repression assay described above was modified to allow independent control of the expression of the two components of the heterotetrameric CC. Specifically, CC-Tet-A was fused to the *C* terminus of a second Lac repressor variant, LacI**, where the dimerization interface was weakened even further with an additional Tyr282→Ala substitution.[49] The resulting protein, LacI**-CC-Tet-A, was expressed from an insulated low-level constitutive promoter, *pro1*.[52] The additional Tyr282→Ala substitution was necessary because LacI* shows dimerization when expressed from this promoter (Figure S8). The regulatory component, CC-Tet-B fused to the *C* terminus of SUMO, was expressed from the arabinose-inducible P_{araBAD} promoter. When this system was tested in *E. coli*, increasing the arabinose concentration, thereby increasing SUMO-CC-Tet-B expression, caused incremental increases in transcription repression (Figures 4b and 4c). The strength of repression did not increase when the inducer concentration was increased above 0.02% arabinose. This was likely because the availability of LacI**-CC-Tet-A monomers became limiting once SUMO-CC-Tet-B reached a certain cellular concentration. As above, the SUMO domain was used here to improve the in-cell properties of CC-Tet-B. However, this tunable repression was also achieved when the peptide was expressed without the SUMO fusion partner, showing that isolated CC peptides can also be used effectively in cells (Figure S9). Altogether, this demonstrates that the use of heterotetramers allows control over complex assembly as designed.

As a further demonstration of the control achievable with these straightforward components, a system where transcription repression could be relieved was designed using the homodimer CC-Di (Figure 4d). This homodimeric CC was fused to the *C* terminus of LacI** and expressed from *pro1*. A second copy of CC-Di was fused to the *C* terminus of SUMO and this protein, SUMO-CC-Di, was expressed from P_{araBAD}. In the absence of SUMO-CC-Di, LacI**-CC-Di was expected to homodimerize and cause transcription repression. On induction with arabinose, SUMO-CC-Di was expected to disrupt these homodimers by sequestering the LacI**-CC-Di monomers in LacI**-CC-Di/SUMO-CC-Di heterodimers, thereby relieving transcription repression. When this system was tested in *E. coli*, strong repression was observed in the absence of arabinose due to LacI**-CC-Di homodimerizing (Figure 4e). In the presence of 0.2% arabinose, however, transcription repression was greatly reduced, indicating that most of the LacI**-CC-Di monomers were sequestered in complexes with the excess SUMO-CC-Di

monomers and were therefore not available to form DNA-binding complexes. Therefore, this system allows transcription repression to be switched off by adding arabinose to cells.

CONCLUSIONS:

The design and characterization of coiled-coil (CC) protein-protein-interaction (PPI) domains for use *in vitro* is a well-established practice.[20,21] We and others are now taking steps to utilize such CCs in cells.[2,28,29,36-38] Here, we present a set of *de novo* designed CCs that function faithfully both *in vitro* and in *E. coli*. We anticipate that these *de novo* CCs will be useful as new PPI domains in a wide range of synthetic-biology applications that require fine control over the spatial and temporal assembly of protein domains in cells. These may include the assembly of proteinaceous nanostructures, the co-localization of enzymes in cascades, and the localization of components to specific sub-cellular environments.

The set covers a broad region of CC structural space and incudes dimeric and tetrameric oligomeric states, parallel and antiparallel topologies, and homo- and heteromeric assemblies. Specifically, the set contains three homomeric CCs that had been characterized previously *in vitro*, and a newly designed heterotetramer, CC-Tet-A_2B_2, which has been designed to fill a hitherto unmet set of requirements for controlling more-complex assemblies. This heterotetramer is achieved by designing an overall acidic and an overall basic peptide, a strategy that has previously been used successfully to generate heteromeric coiled coils.[23,25,40,41,43] When charged residues are sufficiently close to the inter-helical interface, charge repulsion ensures that the peptides do not form homomers. However, these peptides are able to form heteromeric assemblies, where charge complementation relieves the repulsion. Thus, CC-Tet-A_2B_2 is an obligate heterotetramer and adds considerable utility to the *de novo* CC toolkit. In addition to the regulatory application demonstrated here, it will be useful in applications where multiple functionally distinct components must be co-localized and where off-target homomeric interactions are undesirable.

As demonstrated using a transcription repression assay in which the *de novo* CCs replace the oligomerization domain of the Lac repressor, all tested CCs assemble in *E. coli*. Thus, these short and straightforward modules are able to find their interaction partners in a complex cellular environment with no obvious adverse effects on cell viability. However, some significant challenges remain before the design of CCs for *in vivo* applications can be regarded as routine. The CCs are all highly stable *in vitro* and fold in the *E. coli* cytosol, but they are not yet necessarily optimized for performance in cells: their sequences are idealized and highly repetitive and their structures contain high localized charge. Given the prevalence of CC sequences in nature, as well as the charged nature of some of the components, it is possible that the *de novo* peptides may have unwanted cross reactions with endogenous proteins or other macromolecules;[53,54] though we see no significant inhibition of cell growth when our constructs are expressed. Further investigation is required to determine whether the CCs interact with the *E. coli* proteome. Additionally, it will be interesting to determine how predictably the designs behave in the more complex environment of a eukaryotic cell.[55]

Finally, by controlling the CC components at the transcriptional level, we have demonstrated control over complex assembly and function. For some applications it would be useful to achieve faster switching by controlling complex assembly and disassembly post-translationally, for example through small-molecule binding, chemical modification, or, as demonstrated recently, with proteolysis.[38] As tetrameric CCs have relatively large hydrophobic cores, and consequently are generally very stable, they may be particularly tolerant of changes intended to introduce more-sophisticated properties into them. Future work will focus on expanding this set of constitutive CC domains to include inducible interactions.

MATERIALS AND METHODS:

De novo heterotetramer design

The peptides CC-Tet-A and CC-Tet-B contained *a*=Leu/*d*=Ile cores. Glu or Lys residues were placed at all *e* and *g* positions in CC-Tet-A and CC-Tet-B, respectively. In both peptides, Ala was placed at all *b* and *c* positions. *f* positions were populated with Lys residues in CC-Tet-A or Gln in CC-Tet-B to promote solubility and aid purification, and a single Trp residue was placed at position *f*19 to provide a chromophore for concentration determination. Helix-capping Gly residues were added to *N* and *C* termini.

Solid-phase peptide synthesis and peptide purification

All reagents were supplied by Cambridge Reagents or Acros Organics unless otherwise stated. Peptides were synthesized on a 0.1 mmol scale by solid-phase peptide synthesis on Rink amide resin using a Liberty Blue automated microwave peptide synthesizer (CEM) and Fmoc-protected amino acids at 0.2 M in dimethylformamide (DMF). Deprotection was performed with 20% (v/v) morpholine (Alfa Aesar) in DMF. Coupling was performed using 0.5 M 6-Chloro-1-hydroxybenzotriazole in DMF as the activator and 1.0 M N,N'-Diisopropylcarbodiimide in DMF as the activator base. Following synthesis, peptides were *N*-terminally acetylated using acetic anhydride and pyridine in 1:1 DMF/dichloromethane (Sigma Aldrich). Peptides were simultaneously cleaved from the resin and side chain deprotected using trifluoroacetic acid (TFA) with 5% (v/v) H_2O and 5% (v/v) triisopropylsaline. Resin was removed by filtration and crude peptides were precipitated with diethyl ether (Honeywell Research Chemicals) then isolated by centrifugation. Peptides were dissolved in 1:1 ultrapure Milli-Q water/acetonitrile then lyophilized.

Peptides were purified by reversed-phase HPLC using C18 reversed-phase columns (150 x 10 mm, 100 Å pore size, Phenomenex). Linear gradients of ultrapure Milli-Q water containing 0.1% (v/v) TFA and acetonitrile containing 0.1% (v/v) TFA were used. Peptide masses were confirmed by MALDI-TOF mass spectrometry with an ultrafleXtreme II mass spectrometer in positive-ion reflector mode (Bruker, UK) using an α-cyano-4-hydroxycinnamic acid matrix (Fluka Analytical). Fraction purity was confirmed by reversed-phase HPLC using analytical C18 reversed-phase columns (100 x 4.6 mm, 100 Å pore size, Phenomenex) with linear gradients of the above buffers. Selected fractions were pooled and lyophilized prior to biophysical characterization.

Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy was performed using either a JASCO J-810 or a JASCO J-815 spectropolarimeter with a Peltier temperature controller (Jasco, UK). All measurements were performed in phosphate buffered saline (8.2 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4). Samples were analyzed in quartz cuvettes. Full spectra were measured between 190 and 260 nm with a 1 nm step size, 100 $nm.min^{-1}$ scanning speed, 1 nm bandwidth and 1 second response time. Spectra were measured at 5 °C. Variable temperature experiments were performed by heating and cooling samples 5–95–5 °C at a rate of 40 $°C.h^{-1}$ whilst monitoring CD at 222 nm at 0.5 °C intervals. Data was buffer subtracted then CD (mdeg) was converted to mean residue ellipticity (MRE, $deg.cm^2.dmol^{-1}.res^{-1}$) by normalizing for peptide concentration, number of amide bonds and cuvette pathlength. Fraction helix (%) was calculated using the following equation (where MRE_{coil} = 640-45T; T, temperature (°C); n, number of amide bonds in sample including *C*-terminal amide):[56]

$$\text{Fraction helix (\%)} = 100 \times \frac{MRE_{222} - MRE_{coil}}{-42500 \times \left(1 - {}^{3}/_{n}\right) - MRE_{coil}}$$

Sedimentation velocity analytical ultracentrifugation

Sedimentation velocity analysis was conducted on CC-Tet-A_2B_2 at 20 °C in a Beckman-Optima XL-A analytical ultracentrifuge using an An-60 Ti rotor. Solutions were prepared in buffer at a total peptide concentration of 140 μM. Samples were prepared in phosphate buffered saline at a volume of 305 μL and loaded into an SV cell with a 12 mm graphite-filled centerpiece and quartz windows. The reference channel was loaded with 320 μL buffer. Following temperature equilibration at 3 krpm, the sample was spun at 50 krpm. A total of 120 absorbance scans at 280 nm over a radial range of 5.8 to 7.3 cm were measured at 5 min intervals. Data were fitted to a continuous c(s) distribution model using Sedfit, at 99% confidence level.[57] The baseline, meniscus, frictional ratio (f/f_0), and systematic time-invariant and radial-invariant noise were fitted. The partial specific volume ($\bar{v}$) for each peptide/peptide combination was calculated using Sedfit. The buffer density was calculated using SEDNTERP. Example data and fits, continuous c(s) distribution and residuals were plotted with residuals shown as a rectangular greyscale bitmap displaying residuals for every scan, stacked from top to bottom.

Sedimentation equilibrium analytical ultracentrifugation

Sedimentation equilibrium experiments were conducted on heteromer CC-Tet-A_2B_2 using 12 mm six-channel epon-charcoal equilibrium cells with quartz windows and an An-50 rotor and on peptides CC-Tet-A and CC-Tet-B using SV cells with 12 mm graphite-filled centerpieces and quartz windows and an An-60 rotor. All experiments were performed at 20 °C in a Beckman-Optima XL-I or XL-A analytical ultracentrifuge. Solutions were prepared in phosphate buffered saline at a total peptide concentration of 70 μM. CC-Tet-A_2B_2 was spun at 24–39 krpm and the individual peptides were spun at 45–60 krpm. Absorbance scans at 280 nm were measured at 3 krpm intervals following an initial equilibration period of 8 h and a second equilibration period of 1 h. Data sets were initially fitted to a single ideal species model using Ultrascan II (http://www.ultrascan.uthscsa.edu/). 99% confidence limits were calculated using Monte Carlo analysis of the obtained fits.

Plasmid construction

Design of DNA encoding CC peptides. DNA sequences for CC-Di, CC-Tet, apCC-Tet, CC-Tet-A and CC-Tet-B were manually designed to reduce repetitive sequences and to optimize for *E. coli* codon usage (Table S4). DNA sequences also contained 5′ and 3′ recognition sites for restriction enzymes XbaI and Acc65I, respectively. DNA fragments encoding CC peptides were synthesized as GeneArt Strings (Thermo Fisher Scientific) or as GeneStrands (Eurofins Genomics).

GFP reporter plasmid (Kan^R). The reporter plasmid, pVRb-*lacUV5*, contained superfolder green fluorescent protein (sfGFP)[58] expressed from the *lacUV5* promoter. The pVRb-*lacO1-lacO1* reporter plasmid contains a second O_1 *lac* operator spaced 92 bp from the first. Both plasmids have been described previously.[29]

pBAD-LacI (Amp^R) and pVRc-LacI* (Cm^R) plasmids.* The plasmids pBAD-LacI*-CC-Di, pBAD-LacI*-CC-Tet, pBAD-LacI*-apCC-Tet, pBAD-LacI*-CC-Tet-B and pVRc-LacI*-CC-Tet-A were prepared by inserting DNA fragments encoding the peptides into pBAD-LacI* and pVRc-LacI*. The pBAD-LacI* and pVRc-LacI* plasmids have been described in detail previously.[29] These plasmids carried a truncated version of the Lac repressor (residues 1–332) lacking its *C*-terminal tetramerization domain (residues 333–360) and contained a point mutation, L251A, and XbaI and Acc65I restriction sites for the introduction of the peptide-encoding DNA fragments at the Lac repressor *C* terminus *via* a flexible linker (Figure S10).

The LacI*-peptide constructs had *N*-terminal 6His, T7 and Xpress tags and were expressed from the P_{araBAD} arabinose-inducible promoter.

pBAD-6H (Amp^R) plasmids. The plasmid pBAD-6H was prepared using deletion PCR to remove the Lac repressor gene and T7 and Xpress tags from the plasmid pBAD-LacI*. The plasmid pBAD-6H-CC-Tet-B was prepared in the same way from pBAD-LacI*-CC-Tet-B.

pBAD-SUMO (Amp^R) plasmids. The plasmids pBAD-SUMO and pBAD-SUMO-CC-Tet-B were prepared by inserting the gene for yeast SUMO (Smt3, small ubiquitin-related modifier) into pBAD-6H and pBAD-6H-CC-Tet-B, respectively. pBAD-SUMO-CC-Di was made by inserting the DNA for CC-Di into pBAD-SUMO. In pBAD-SUMO-CC-Di and pBAD-SUMO-CC-Tet-B the peptide genes were fused to the *C* terminus of SUMO *via* a flexible linker (Figure S10). In all plasmids, SUMO was also fused to an *N*-terminal 6 His tag.

Y282A Point mutation. Site-directed mutagenesis was used to introduce a point mutation, Y282A, into the LacI* gene to produce the LacI** variant in pBAD-LacI**-CC-Di, pVRc-LacI**-CC-Tet-A and pBAD-LacI**-WTtet. PCR was performed with mutagenesis primers and the resulting blunt-end products were re-circularized.

Pro1 constitutive promoter. The DNA sequence for *pro1* [52] was flanked with 5′ and 3′ recognition sites for restriction enzymes SpeI and NcoI, respectively (Table S4). The promoter sequence was synthesized as a GeneArt String (Thermo Fisher Scientific).

Constitutive promoter plasmids. The P_{araBAD} promoter and *araC* gene were removed from, and an SpeI restriction site was introduced into plasmids pVRc-LacI**, pVRc-LacI**-CC-Di and pVRc-LacI**-CC-Tet-A. DNA for *pro1* was inserted into these plasmids to produce pVRc-*pro1*-LacI**, pVRc-*pro1*-LacI**-CC-Di and pVRc-*pro1*-LacI**-CC-Tet-A.

Transcription repression assays

GFP transcription repression assays were performed as previously described.[29] TB28 (MG1655ΔLacIZYA) *E. coli* [59] were transformed with the reporter plasmid and sample plasmids carrying the proteins of interest. Following antibiotic selection on LB agar plates, three or four colonies were selected and grown overnight at 37 °C in 5 mL supplemented M9 media (M9 minimal media: 42.3 mM Na_2HPO_4; 21.6 mM KH_2PO_4; 18.7 mM NH_4Cl; 8.56 mM NaCl. Supplements: 0.2% (w/v) casamino acids; 10 mM $CaCl_2$; 0.25% (v/v) glycerol; 2 mM $MgSO_4$; 2 μg.mL^{-1} thiamine) with antibiotics as required (ampicillin, 100 μg/mL; chloramphenicol, 25 μg/mL; kanamycin, 50 μg/mL). 10 mL fresh supplemented M9 media was inoculated with overnight culture and induced with varying arabinose concentrations (0.0–0.2% (w/v)). Cultures were grown at 37 °C until OD_{600} ≈ 0.4–0.6. OD_{600} values were recorded to three significant figures using a Lambda Bio UV/Vis spectrometer (PerkinElmer). 5 mL of each culture was pelleted (10 min, 5000 rpm, 4 °C) then re-suspended in 250 μL PBS. Two 100 μL aliquots of each sample were loaded into consecutive wells of a black 96 well polypropylene microplate. GFP fluorescence was measured using a FlexStation Microplate Reader (Molecular Devices) with excitation and emission wavelengths of 470 nm and 510 nm, respectively. GFP/OD_{600} was calculated for each 100 μL sample then the duplicate measurements were averaged. Next, the technical replicates were averaged. Mean values were plotted on a bar chart with error bars as one SD from the mean. Fold-repression values were calculated relative to the GFP-only control unless otherwise stated.

ACKNOWLEDGEMENTS:

CLE was supported by the EPSRC and BBSRC Centre for Doctoral Training in Synthetic Biology (EP/L016494/1) and a BBSRC grant awarded to NJS and DNW (BB/S002820/1). AJS was supported by BrisSynBio, a BBSRC/EPSRC Synthetic Biology Research Centre (BB/L01386X/1). JLB was supported by a European Research Council Advanced Grant awarded to DNW (340764). DNW held a Royal Society Wolfson Research Merit Award (WM140008). We also thank the University of Bristol School of Chemistry Mass Spectrometry Facility for use of the EPSRC-funded Bruker Ultraflex MALDI-TOF/TOF instrument (EP/K03927X/1).

AUTHOR CONTRIBUTIONS:

CLE and AJS designed and performed the experiments. JLB contributed to protein design. NJS and DNW conceived the study and contributed to experimental design. CLE, NJS and DNW wrote the manuscript.

NOTES:

The authors declare no competing financial interest.

SUPPORTING INFORMATION

Supplementary figures S1–S10, supplementary tables S1–S4 and supplementary references.

REFERENCES:

(1) Nooren, I. M. A., and Thornton, J. M. (2003) Diversity of protein-protein interactions, *EMBO J. 22*, 3486-3492.

(2) Lee, M. J., Mantell, J., Brown, I. R., Fletcher, J. M., Verkade, P., Pickersgill, R. W., Woolfson, D. N., Frank, S., and Warren, M. J. (2018) De novo targeting to the cytoplasmic and luminal side of bacterial microcompartments, *Nat. Commun. 9*.

(3) Lee, M. J., Mantell, J., Hodgson, L., Alibhai, D., Fletcher, J. M., Brown, I. R., Frank, S., Xue, W. F., Verkade, P., Woolfson, D. N., et al. (2018) Engineered synthetic scaffolds for organizing proteins within the bacterial cytoplasm, *Nat. Chem. Biol. 14*, 142-147.

(4) Sciore, A., Su, M., Koldewey, P., Eschweiler, J. D., Diffley, K. A., Linhares, B. M., Ruotolo, B. T., Bardwell, J. C. A., Skiniotis, G., and Marsh, E. N. G. (2016) Flexible, symmetry-directed approach to assembling protein cages, *Proc. Natl. Acad. Sci. U.S.A. 113*, 8681-8686.

(5) Badieyan, S., Sciore, A., Eschweiler, J. D., Koldewey, P., Cristie-David, A. S., Ruotolo, B. T., Bardwell, J. C. A., Su, M., and Marsh, E. N. G. (2017) Symmetry-Directed Self-Assembly of a Tetrahedral Protein Cage Mediated by de Novo-Designed Coiled Coils, *Chembiochem 18*, 1888-1892.

(6) Brophy, J. A. N., and Voigt, C. A. (2014) Principles of genetic circuit design, *Nat. Methods 11*, 508-520.

(7) Lloyd, G., Landini, P., and Busby, S. (2001) Activation and repression of transcription initiation in bacteria, *Essays Biochem. 37*, 17-31.

(8) Browning, D. F., and Busby, S. J. W. (2004) The regulation of bacterial transcription initiation, *Nat. Rev. Microbiol. 2*, 57-65.

(9) Dove, S. L., and Hochschild, A. (1998) Conversion of the omega subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target, *Genes Dev. 12*, 745-754.

(10) Joung, J. K., Ramm, E. I., and Pabo, C. O. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions, *Proc. Natl. Acad. Sci. U.S.A. 97*, 7382-7387.

(11) Pomerantz, J. L., Wolfe, S. A., and Pabo, C. O. (1998) Structure-based design of a dimeric zinc finger protein, *Biochemistry 37*, 965-970.

(12) Hou, J. R., Zeng, W. Q., Zong, Y. Q., Chen, Z. H., Miao, C. S., Wang, B. J., and Lou, C. B. (2018) Engineering the Ultrasensitive Transcription Factors by Fusing a Modular Oligomerization Domain, *ACS Synth. Biol. 7*, 1188-1194.

(13) Younger, A. K. D., Dalvie, N. C., Rottinghaus, A. G., and Leonard, J. N. (2017) Engineering Modular Biosensors to Confer Metabolite-Responsive Regulation of Transcription, *ACS Synth. Biol. 6*, 311-325.

(14) Becker, N. A., Schwab, T. L., Clark, K. J., and Maher, L. J. (2018) Bacterial gene control by DNA looping using engineered dimeric transcription activator like effector (TALE) proteins, *Nucleic Acids Res. 46*, 2690-2696.

(15) Chang, H. J., Mayonove, P., Zavala, A., De Visch, A., Minard, P., Cohen-Gonsaud, M., and Bonnet, J. (2018) A Modular Receptor Platform To Expand the Sensing Repertoire of Bacteria, *ACS Synth. Biol. 7*, 166-175.

(16) Stites, W. E. (1997) Protein-protein interactions: Interface structure, binding thermodynamics, and mutational analysis, *Chem. Rev. 97*, 1233-1250.

(17) O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) X-Ray Structure of the Gcn4 Leucine Zipper, a 2-Stranded, Parallel Coiled Coil, *Science 254*, 539-544.

(18) Alberti, S., Oehler, S., Vonwilckenbergmann, B., and Mullerhill, B. (1993) Genetic-Analysis of the Leucine Heptad Repeats of Lac Repressor - Evidence for a 4-Helical Bundle, *EMBO J. 12*, 3227-3236.

(19) Glover, J. N. M., and Harrison, S. C. (1995) Crystal-Structure of the Heterodimeric Bzip Transcription Factor C-Fos-C-Jun Bound to DNA, *Nature 373*, 257-261.

(20) Woolfson, D. N. (2005) The design of coiled-coil structures and assemblies, *Adv. Protein Chem. 70*, 79-112.
(21) Woolfson, D. N. (2017) Coiled-Coil Design: Updated and Upgraded, In *Fibrous Proteins: Structures and Mechanisms* (Parry, D., Squire, J., Ed.), pp 35-61, Springer International Publishing AG.
(22) Lupas, A. N., and Gruber, M. (2005) The structure of alpha-helical coiled coils, *Adv. Protein Chem. 70*, 37-78.
(23) Thomas, F., Boyle, A. L., Burton, A. J., and Woolfson, D. N. (2013) A Set of de Novo Designed Parallel Heterodimeric Coiled Coils with Quantified Dissociation Constants in the Micromolar to Sub-nanomolar Regime, *J. Am. Chem. Soc. 135*, 5161-5166.
(24) Fletcher, J. M., Boyle, A. L., Bruning, M., Bartlett, G. J., Vincent, T. L., Zaccai, N. R., Armstrong, C. T., Bromley, E. H. C., Booth, P. J., Brady, R. L., et al. (2012) A Basis Set of de Novo Coiled-Coil Peptide Oligomers for Rational Protein Design and Synthetic Biology, *ACS Synth. Biol. 1*, 240-250.
(25) O'Shea, E. K., Lumb, K. J., and Kim, P. S. (1993) Peptide Velcro - Design of a Heterodimeric Coiled-Coil, *Curr. Biol. 3*, 658-667.
(26) Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M., and Hodges, R. S. (1996) Formation of parallel and antiparallel coiled-coils controlled by the relative positions of alanine residues in the hydrophobic core, *J. Biol. Chem. 271*, 3995-4001.
(27) Rhys, G. G., Wood, C. W., Beesley, J. L., Zaccai, N. R., Burton, A. J., Brady, R. L., Thomson, A. R., and Woolfson, D. N. (2019) Navigating the Structural Landscape of De Novo alpha-Helical Bundles, *J. Am. Chem. Soc. 141*, 8787-8797.
(28) Hussey, B. J., and McMillen, D. R. (2018) Programmable T7-based synthetic transcription factors, *Nucleic Acids Res. 46*, 9842-9854.
(29) Smith, A. J., Thomas, F., Shoemark, D., Woolfson, D. N., and Savery, N. J. (2019) Guiding Biomolecular Interactions in Cells Using de Novo Protein-Protein Interfaces, *ACS Synth. Biol. 8*, 1284-1293.
(30) Wolfe, S. A., Ramm, E. I., and Pabo, C. O. (2000) Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers, *Struct. Fold. Des. 8*, 739-750.
(31) Wolfe, S. A., Grant, R. A., and Paboll, C. O. (2003) Structure of a designed dimeric zinc finger protein bound to DNA, *Biochemistry 42*, 13401-13409.
(32) Thompson, K. E., Bashor, C. J., Lim, W. A., and Keating, A. E. (2012) SYNZIP Protein Interaction Toolbox: in Vitro and in Vivo Specifications of Heterospecific Coiled-Coil Interaction Domains, *ACS Synth. Biol. 1*, 118-129.
(33) Pelletier, J. N., Campbell-Valois, F. X., and Michnick, S. W. (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments, *Proc. Natl. Acad. Sci. U.S.A. 95*, 12141-12146.
(34) Ghosh, I., Hamilton, A. D., and Regan, L. (2000) Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein, *J. Am. Chem. Soc. 122*, 5658-5659.
(35) Hu, C. D., Chinenov, Y., and Kerppola, T. K. (2002) Visualization of interactions among bZip and Rel family proteins in living cells using bimolecular fluorescence complementation, *Mol. Cell 9*, 789-798.
(36) Pack, P., Muller, K., Zahn, R., and Pluckthun, A. (1995) Tetravalent Miniantibodies with High Avidity Assembling in Escherichia-Coli, *J. Mol. Biol. 246*, 28-34.
(37) Swainsbury, D. J. K., Harniman, R. L., Di Bartolo, N. D., Liu, J. T., Harper, W. F. M., Corrie, A. S., and Jones, M. R. (2016) Directed assembly of defined oligomeric photosynthetic reaction centres through adaptation with programmable extra-membrane coiled-coil interfaces, *Biochim. Biophys. Acta, Bioenerg. 1857*, 1829-1839.
(38) Fink, T., Lonzaric, J., Praznik, A., Plaper, T., Merljak, E., Leben, K., Jerala, N., Lebar, T., Strmsek, Z., Lapenta, F., et al. (2019) Design of fast proteolysis-based signaling and logic circuits in mammalian cells, *Nat. Chem. Biol. 15*, 115-122.

(39) Zaccai, N. R., Chi, B., Thomson, A. R., Boyle, A. L., Bartlett, G. J., Bruning, M., Linden, N., Sessions, R. B., Booth, P. J., Brady, R. L., et al. (2011) A de novo peptide hexamer with a mutable channel, *Nat. Chem. Biol. 7*, 935-941.
(40) Vu, C., Robblee, J., Werner, K. M., and Fairman, R. (2001) Effects of charged amino acids at b and c heptad positions on specificity and stability of four-chain coiled coils, *Protein Sci. 10*, 631-637.
(41) Root, B. C., Pellegrino, L. D., Crawford, E. D., Kokona, B., and Fairman, R. (2009) Design of a heterotetrameric coiled coil, *Protein Sci. 18*, 329-336.
(42) Fairman, R., Chao, H. G., Mueller, L., Lavoie, T. B., Shen, L. Y., Novotny, J., and Matsueda, G. R. (1995) Characterization of a New 4-Chain Coiled-Coil - Influence of Chain-Length on Stability, *Protein Sci. 4*, 1457-1469.
(43) Fairman, R., Chao, H. G., Lavoie, T. B., Villafranca, J. J., Matsueda, G. R., and Novotny, J. (1996) Design of heterotetrameric coiled coils: Evidence for increased stabilization by Glu(-)-Lys(+) ion pair interactions, *Biochemistry 35*, 2824-2829.
(44) Deng, Y. Q., Liu, J., Zheng, Q., Li, Q. N., Kallenbach, N. R., and Lu, M. (2008) A Heterospecific Leucine Zipper Tetramer, *Chem. Biol. 15*, 908-919.
(45) Cherry, J. L., and Adler, F. R. (2000) How to make a Biological Switch, *J. Theor. Biol. 203*, 117-133.
(46) Bashor, C. J., Patel, N., Choubey, S., Beyzavi, A., Kondev, J., Collins, J. J., and Khalil, A. S. (2019) Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies, *Science 364*, 593-597.
(47) Chen, J., and Matthews, K. S. (1992) Deletion of Lactose Repressor Carboxyl-Terminal Domain Affects Tetramer Formation, *J. Biol. Chem. 267*, 13843-13850.
(48) Dong, F. M., Spott, S., Zimmermann, O., Kisters-Woike, B., Muller-Hill, B., and Barker, A. (1999) Dimerisation mutants of Lac repressor. I. A monomeric mutant, L251A, that binds lac operator DNA as a dimer, *J. Mol. Biol. 290*, 653-666.
(49) Chakerian, A. E., and Matthews, K. S. (1991) Characterization of Mutations in Oligomerization Domain of Lac Repressor Protein, *J. Biol. Chem. 266*, 22206-22214.
(50) Friedman, A. M., Fischmann, T. O., and Steitz, T. A. (1995) Crystal-Structure of Lac Repressor Core Tetramer and Its Implications for DNA Looping, *Science 268*, 1721-1727.
(51) Becker, N. A., Peters, J. P., Lionberger, T. A., and Maher, L. J. (2013) Mechanism of promoter repression by Lac repressor-DNA loops, *Nucleic Acids Res. 41*, 156-166.
(52) Davis, J. H., Rubin, A. J., and Sauer, R. T. (2011) Design, construction and characterization of a set of insulated bacterial promoters, *Nucleic Acids Res. 39*, 1131-1141.
(53) Schavemaker, P. E., Smigiel, W. M., and Poolman, B. (2017) Ribosome surface properties may impose limits on the nature of the cytoplasmic proteome, *Elife 6*.
(54) Mu, X., Choi, S., Lang, L., Mowray, D., Dokholyan, N. V., Danielsson, J., and Oliveberg, M. (2017) Physicochemical code for quinary protein interactions in Escherichia coli, *Proc. Natl. Acad. Sci. U.S.A. 114*, E4556-E4563.
(55) Lebar, T., Lainscek, D., Merljak, E., Aupic, J., and Jerala, R. (2020) A tunable orthogonal coiled-coil interaction toolbox for engineering mammalian cells, *Nat. Chem. Biol. 6*.
(56) Myers, J. K., Pace, C. N., and Scholtz, J. M. (1997) A direct comparison of helix propensity in proteins and peptides, *Proc. Natl. Acad. Sci. U.S.A. 94*, 2833-2837.
(57) Brown, P. H., and Schuck, P. (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation, *Biophys. J. 90*, 4651-4661.
(58) Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein, *Nat. Biotechnol. 24*, 79-88.
(59) Bernhardt, T. G., and de Boer, P. A. J. (2003) The Escherichia coli amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway, *Mol. Microbiol. 48*, 1171-1182.

FIGURES:

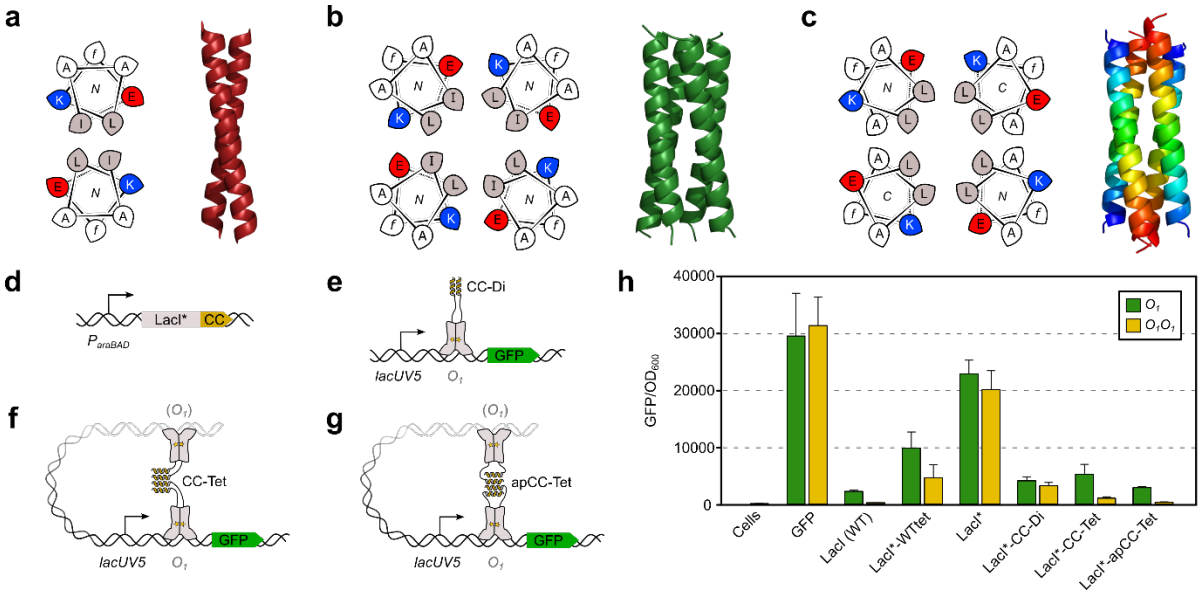


Figure 1 *De novo* designed homomeric coiled coils direct protein-protein interactions in *E. coli*. Helical wheels and X-ray crystal structures for the designed CCs (a) CC-Di,[24] (b) CC-Tet [39] and (c) apCC-Tet.[27] (d) Schematic of the transcription repression assay using a mutant of the Lac repressor, LacI*, showing the general architecture of the LacI*-CC fusion genes. Schematic of the transcription repression assay showing cartoons of (e) a LacI*-CC-Di dimer bound to a single DNA site, (f) a LacI*-CC-Tet tetramer bound to two DNA sites and (g) a LacI*-apCC-Tet tetramer bound to two DNA sites. The reporter gene, GFP, was expressed from the *lacUV5* promoter with or without an additional O_1 *lac* operator placed 92 bp upstream of the *lacUV5* O_1 operator. (h) Transcription repression assay results for LacI*-CC-Di, LacI*-CC-Tet and LacI*-apCC-Tet with each protein expressed at basal levels from P_{araBAD} and with one or two O_1 *lac* operator sites in the reporter plasmid. Error bars are one SD from the mean, N=3.

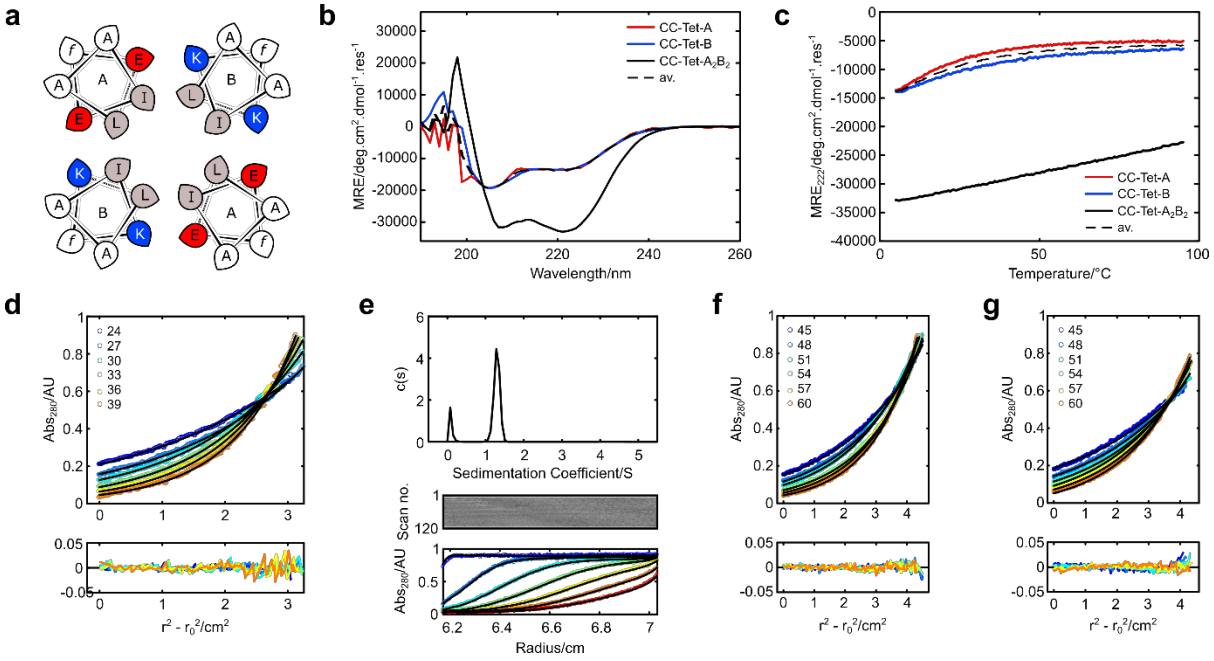


Figure 2 A *de novo* designed heterotetrameric coiled coil, CC-Tet-A_2B_2. (a) Helical wheel for the designed CC, CC-Tet-A_2B_2. (b) Representative CD spectra at 5 °C and (c) variable temperature CD measurements for the peptides CC-Tet-A and CC-Tet-B and the heterotetramer CC-Tet-A_2B_2, monitoring mean residue ellipticity at 222 nm (MRE_{222}) between 5 and 95 °C. (d) Sedimentation equilibrium and (e) sedimentation velocity analytical ultracentrifugation data for the heterotetramer CC-Tet-A_2B_2 returning molecular weights of 12772 Da (3.9 x mean monomer mass) and 13190 Da (4.1 x mean monomer mass), respectively. Sedimentation equilibrium analytical ultracentrifugation data for (f) CC-Tet-A and (g) CC-Tet-B returning molecular weights of 3591 Da (1.1 x monomer mass) and 3153 Da (1.0 x monomer mass), respectively. All measurements were performed in phosphate buffered saline (pH 7.4).

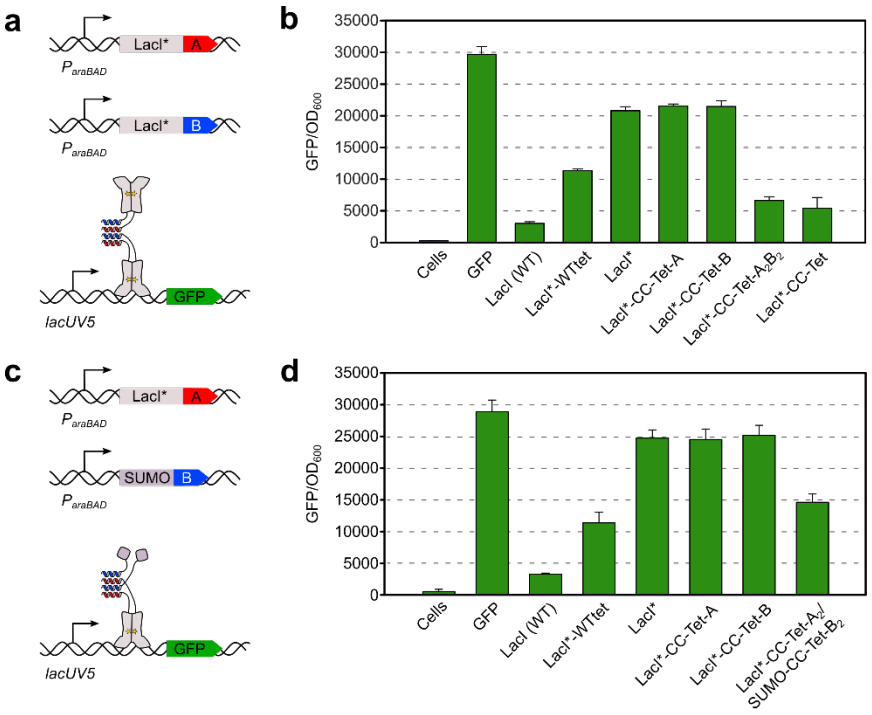


Figure 3 The *de novo* designed heterotetrameric coiled coil folds in *E. coli*. (a) Schematic of the transcription repression assay showing the general architecture of the LacI*-CC fusion genes and a LacI*-CC-Tet-A_2B_2 tetramer bound to a single DNA site. (b) Transcription repression assay results for LacI*-CC-Tet-A_2B_2 with both proteins expressed at basal levels from P_{araBAD} and with a single O_1 *lac* operator site in the reporter plasmid. Error bars are one SD from the mean, N=4. (c) Schematic of the transcription repression assay showing the general architecture of the LacI*-CC-Tet-A and SUMO-CC-Tet-B fusion genes and a LacI*-CC-Tet-A_2/SUMO-CC-Tet-B_2 tetramer bound to a single DNA site. (d) Transcription repression assay results for LacI*-CC-Tet-A_2/SUMO-CC-Tet-B_2 with both proteins expressed at basal levels from P_{araBAD} and with a single O_1 *lac* operator site in the reporter plasmid. Error bars are one SD from the mean, N=4.

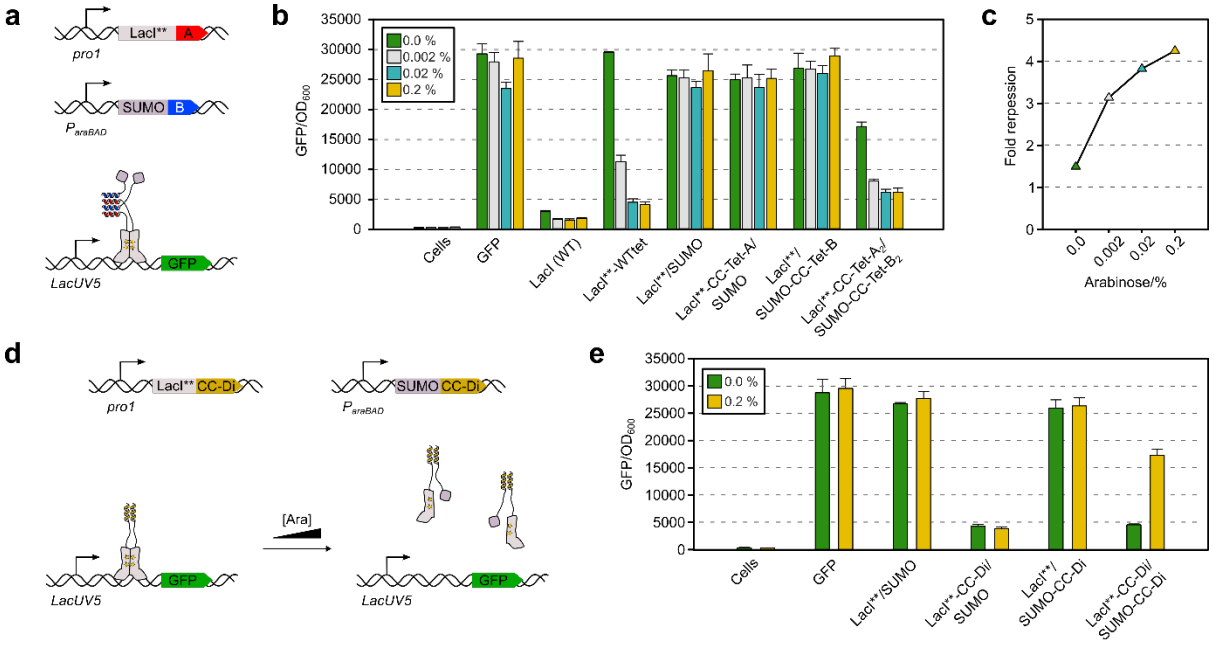


Figure 4 Control of protein-protein interactions in *E. coli*. (a) Schematic of the transcription repression assay using LacI** showing the architecture of the LacI**-CC and SUMO-CC fusion genes and a LacI**-CC-Tet-A_2/SUMO-CC-Tet-B_2 tetramer bound to a single DNA site. (b) Transcription repression assay results and (c) fold repression values for LacI**-CC-Tet-A_2/SUMO-CC-Tet-B_2 with LacI**-CC-Tet-A expressed from *pro1* and SUMO-CC-Tet-B expressed from P_{araBAD} with and without induction of SUMO-CC-Tet-B by various arabinose concentrations (0.0–0.2%) and with a single O_1 *lac* operator site in the reporter plasmid. Fold repression values were calculated relative to the LacI**/SUMO no-coiled coil control. Error bars are one SD from the mean, *N*=4. (d) Schematic of the transcription repression assay using LacI** showing the architecture of the LacI**-CC-Di and SUMO-CC-Di fusion genes (top), a LacI**-CC-Di homodimer bound to a single DNA site (bottom, left) and LacI**-CC-Di/SUMO-CC-Di heterodimers displaced from DNA (bottom, right). (e) Transcription repression assay results for LacI**-CC-Di expressed from *pro1* in the presence of SUMO-CC-Di expressed from P_{araBAD} with and without induction of SUMO-CC-Di by 0.2% arabinose and with a single O_1 *lac* operator site in the reporter plasmid. Error bars are one SD from the mean, *N*=4.

TABLES:

Table 1 Sequences of all discussed peptides.

Peptide	Sequence	Reference
Register*	g abcdefg abcdefg abcdefg abcdef	
CC-Di†	Ac – G E IAALKQE IAALKKE NAALKWE IAALKQ GYY – NH_2	[24]
CC-Tet‡	Ac – G E LAAIKQE LAAIKKE LAAIKWE LAAIKQ GAG – NH_2	[39]
apCC-Tet	Ac - G E LEALAQE LEALAKK LKALAWK LKALAQ G - NH_2	[27]
Register*	cdefg abcdefg abcdefg abcdefg ab	
CC-Tet-A	Ac – G AIEKE LAAIEKE LAAIEWE LAAIEKE LA G – NH_2	N.A.
CC-Tet-B	Ac – G AIKQK LAAIKQK LAAIKWK LAAIKQK LA G – NH_2	N.A.

All peptides are *N*-terminally acetylated and *C*-terminally amidated.

* Peptide heptad registers are indicated above sequences. Peptides are in *c*- or *g*-register (*i.e.*, excluding the capping Gly residues, they begin at a *c* or *g* heptad position).

† YY mass tag was not included in CC-Di sequence used in *E. coli*.

‡ AG mass tag was not included in CC-Tet sequence used in *E. coli*.